

# 청각장애아동의 구화교육을 위한 애니메이션 콘텐츠 제작기술 개발

조성업\* · 오범수\*\* · 임철수\*\*\*

## 1. 개요

구화교육이란 청각장애를 가진 이가 화자를 상대로 해서 “독화(Speech Reading) + 제스처 + 표정 + 잔청능력” 등을 취합하여 화자의 이야기를 이해할 수 있는 능력을 최대한 향상시키는 것을 목표로 하는 학습이다. 디지털 기술 및 IT 기술의 발전과 더불어 교육, 학습 분야에서도 e-learning, 멀티미디어 학습 등 다양한 형태의 교육 방법과 자료들이 생겨나고 있지만, 아직까지도 청각 장애를 가진 이들을 위한 콘텐츠의 개발이나 그러한 콘텐츠를 개발하기 위한 기술 개발은 활발히 이루어지지 않고 있는 것이 현실이다. 또한 이들을 대상으로 하는 우리말 구화 학습 툴이 상품화되어 나와 있지만 대부분이 외산 S/W에 기초를 두고 있어 우리말 구화 교육에는 적절치 않다고 할 수 있다. 이러한 의미에서 본 연구는 청각 장애를 가진 아동들에 대한 우리말 발성, 발음 교육을 효율적으로 수행할 수 있도록 하는 우리말 학습 S/W의 개발과, 청각 장애를 가진 아동들과 일반 아동들이 함께 보고 즐길 수 있는 애니메이션 콘텐츠 제작에 유용하게 활용될 수 있는 3D 애니메이션 기술 개발을 최종 목표로 하며 이를 위해서 다음 그

림 1과 같이 3부분으로 나누어 기술 개발 및 최종 결과물 제작을 진행했다.

• 음성 기반 자동 립싱크 애니메이션 기술 개발 : 음성 데이터(wave file)를 입력으로 받아, 음성의 발음 시간을 분석한 후, 발음 시간을 기본으로 해서 자동으로 립싱크 애니메이션 키프레임 정보를 생성해내는 기술이다. 이때 생성된 키프레임 정보를 이용해 입술의 겉모습의 움직임 뿐 아니라, 발성 기관의 내부 구조, 얼굴의 근육 애니메이션도 동시에 자동으로 생성해 낼 수 있게 된다. 자동 립싱크 애니메이션 생성 기술을 기본으로 해서, 자연스러운 얼굴 표정 생성 및 발음 시 호흡 정보, 입술의 움직임 정보 등도 자동으로 생성해 낼 수 있는 기능을 같이 개발하게 되며 이 모든

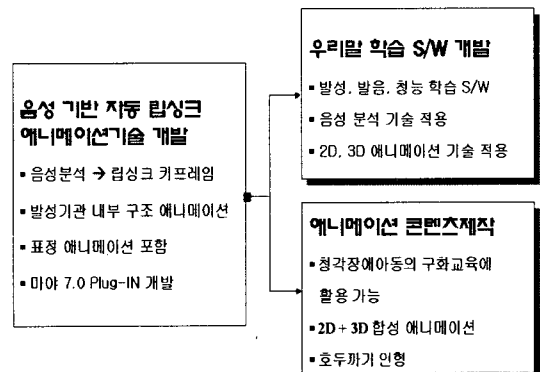


그림 1. 연구 개요

\* 비주텍쓰리디(주) 책임연구원  
 \*\* 종신회원  
 \*\*\* 서경대학교 컴퓨터공학과 교수

기능들은 상용 3D 툴인 Maya 7.0의 Plug-IN으로 통합되어 개발된다.

- **우리말 학습 S/W 개발:** 우리말 학습 S/W는 청각 장애를 가진 아동들을 대상으로 해서 우리말의 발성, 발음을 교육하고, 발음에 대한 청능을 평가할 수 있는 기능을 가지도록 개발되는 툴이다. 여기에서는 음성 분석 기술이 적용되어서 사용자의 발성의 높이, 발성의 크기, 발성의 유무, 발성의 지속 시간, 발성의 시작, 발성에 포함된 유성음과 무성음의 구분 등을 분석해 그 결과를 2D 애니메이션으로 표현해 사용자에게 보여주게 되며, DB에 포함된 발음을 얼마나 정확하게 듣고 따라하는지에 대한 분석과 평가도 같이 이루어져서 청각장애아동의 발음과 청능을 평가하고 역시 2D 애니메이션으로 그 결과를 사용자에게 보여주게 된다. 또한 이러한 발음의 청능 평가에서는 음성 기반 자동 립싱크 애니메이션 기술을 활용한 3D 립싱크 애니메이션 및 발성 기관 내부 구조 애니메이션을 제작하고 DB로 구축해 활용함으로써 우리말 학습의 효과를 극대화 시킬 수 있도록 한다.

- **애니메이션 콘텐츠 제작:** 디지털 기술의 발달과 더불어 엔터테인먼트 분야에서도 아동들이 보고 즐길 수 있는 많은 콘텐츠들이 쏟아져 나오고 있지만, 아직까지는 청각 장애아동들이 쉽게 이해할 수 있는 형식과 구성을 가진 콘텐츠의 제작과 그러한 제작을 뒷받침할 수 있는 기술 개발은 일반적인 추세가 아니다. 본 연구에서는 음성 기반 자동 립싱크 애니메이션 기술을 활용해 우리말을 발음할 때의 실제 입술의 움직임을 시뮬레이션하고 애니메이션 콘텐츠 제작에 이 애니메이션 정보를 활용해 일반 애니메이션에 비해서 청각 장애를 가진 아동들이 주인공들의 발음을

구화를 통해 이해할 수 있도록 하는 콘텐츠를 제작한다. 본 과제에서 선정된 콘텐츠는 “호두까기 인형”을 10분 정도의 분량으로 2D와 3D를 합성해 제작한다.

## 2. 기술 현황

### 2.1 세계기술 현황

언어 학습 툴로써 IBM의 Speech viewer나 Tell-me-more라는 상용 프로그램들이 선보이고 있으며, 음성 발성 기관을 보다 자연스럽게 모델링하려고 연구하고 있다. Speech viewer는 언어 교정을 위해 개발되고 있으며, Tell-me-more는 주로 외국어 학습자를 위해서 다양한 외국어를 학습 서비스와 병행으로 개발되고 있다.

조음 음성 합성기에서 성도의 모델링에 대한 연구로 실제적으로 인간의 성도 구조와 유사한 조음 음성 합성기를 만들고 이 합성기를 다양한 크기의 단면적을 갖는 튜브의 결합으로 간주하여 각 튜브의 단면적을 구함으로써 음성의 발생과정에 대한 시뮬레이션을 시도해 왔다. 그러나 아직까지 단일 합성기에 대한 추정은 실제 성도 모양과 많은 차이를 보여 언어학습에 적합하지 않다.

음성 분석 엔진 및 그를 이용한 립싱크 애니메이션 적용 기술은 해외에서는 이미 오래전부터 패키지 소프트웨어 및 상용 3D 저작 S/W와 연계하여 애니메이션 제작에 실제로 활용되면서 꾸준히 연구 개발이 진행되고 있는 상황이다.

실제로 애니메이션을 제작하는 경우 우선 성우가 대사 부분을 녹음하고 이에 맞추어 립싱크 애니메이션을 제작하는 방식이 대세를 이루고 있다. 지금까지 제품화 되어 판매되고 있는 “음성 분석 기술”을 이용한 “립싱크 애니메이션” 관련 제품들은 다음과 같다.



그림 2. 자동 립싱크 애니메이션 관련 해외 제품군들

최근에는 음성 데이터를 분석해 립싱크 애니메이션을 직접 수행하는 기술 외에, 입술을 포함한 얼굴 전체의 움직임을 직접 캡처해서 표정 및 립싱크 작업을 동시에 사실적으로 처리하는 Performance Capture 기법이 개발되어 영화 제작 등에 사용되고 있다. Performance Capture에서 배우들은 얼굴과 머리에 150개 정도의 광반사판을 붙이고 표정 및 대사 연기에 임해서 배우들의 표정 및 얼굴 움직임 하나하나를 디지털 카메라에 전달시키면 카메라는 얼굴의 표정과 움직임을 3차원 점들의 연결로 기록한다. 3차원 디지털 이미지로 변환된 배우들의 연기는 이미 만들어놓은 3D 얼굴 모델 위에 입혀지고 자연스러운 얼굴 표정 및 움직임으로 표현된다.

## 2.2 국내 기술 현황

국내에서 주로 사용되는 언어 학습 툴은 영어에 기반을 둔 외국 제품을 주로 사용하고 있으며, 순수 국내 기술을 이용해 우리말을 기반으로 한 국산 구화 교육 전문 학습 툴은 아직 개발되어 있지 않다. 또한, 음성 분석 엔진 및 그를 이용한 립싱크 애니메이션 적용 기술은 패키지 소프트웨어 및 상용 3D 저작 S/W와 연계하여 애니메이션 제작에 실제로 활용된 예는 아직 없는 실정이다.

음성학적 연구는 있으나 언어치료를 목적으로 음성의 시각화에 대해서는 외국 제품에 의존하여

국산화 소프트웨어는 아직 개발되어 있지 않다.

국내에서는 언어치료를 위해 현재 Speech Viewer라는 IBM에서 개발된 시청각 피드백이 가능한 소프트웨어를 가장 많이 사용하고 있는 상태이다. 한글 버전이 출시되었으나 영어를 기반으로 하고 있어 청각장애아동의 수준에 따라 재구성해야 하는 불편과 한글과 영어의 발성 및 발어의 차이 등으로 인하여 우리말 학습 툴로서 적합하지 않은 실정이다.

전문 인력의 부족 및 기술력 부족으로 인해 국내 애니메이션 제작 업체들의 인-하우스 소프트웨어 개발 역량은 전반적으로 부족한 편이지만, 최근 들어 투자의 확대와 인식 변화에 힘입어 국책 연구 과제 및 대형 스튜디오를 중심으로 프로젝트를 진행하면서 활발한 연구개발이 진행되고 있는 중이다.

순수 국내 기술로 국내 기업에서 개발하고 있는 립싱크 제품은 (주)이오리스에서 제작한 한국어 립싱크 플러그인이 있다. 그 외에 한국툰봄, RealTalk, MotionTechnologies 등에서 음성 분석을 이용한 립싱크 애니메이션 기술을 개발한 경우가 있지만 LipSinc사와 같은 외국 개발사의 기술을 들여와 활용한 정도에 그쳤으므로 순수 국내 기술로 개발된 제품이라고 보기는 어렵다.

## 3. 청각장애아동의 구화교육을 위한 애니메이션콘텐츠 제작기술 개발

### 3.1 음성분석을 통한 발성의 시각화 기술

음성분석을 기본으로 해서 다음과 같은 기능을 가지는 라이브러리를 개발했다. 아래 8가지 라이브러리는 모두 시간에 대해서 분석 결과를 수치적 데이터로 분석해 그 결과를 반환한다.

- 발성과 발음의 그래프 시각화 함수

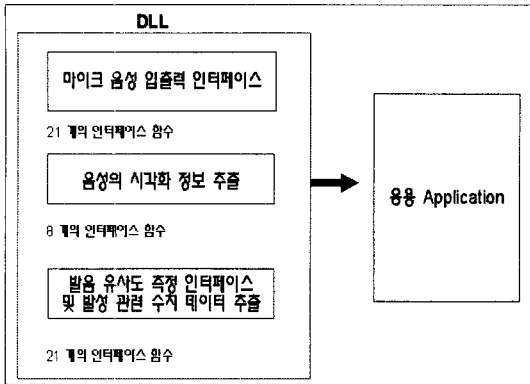


그림 3. 발성/발화 음성 분석 모듈 인터페이스 개요

- 두 발음의 유사도 분석
- 현재 소리가 있는지 없는지 분석
- 소리의 강도(크기) 측정 분석
- 발성이 시작되었는지의 유무를 분석
- 현재 소리가 얼마 동안 지속되고 있는지 분석
- 소리의 높이 측정 분석
- 소리의 유성음/ 무성음 여부 판단 분석

음성 분석 기술에서 두 발음의 유사도를 계산하기 위해서 사용된 알고리즘은 다음과 같다.

▪ **Feature Extraction** : 표준 발음과 발성 화자와의 발음 유사도를 측정하기에 앞서 먼저 입력된 음성 파형의 특징을 검출해내는 feature extraction 과정을 필요로 한다. 여기서 사용된 기법은 Mel-frequency Cepstral Coefficient (MFCC) 방법으로 시간 축 음성 데이터를 주파수 특성을 분석한 뒤에 주파수 값을 잘 구분 짓는 Mel scale로 spectrum 값을 필터링 한다. 그 다음에 인간의 귀가 소리를 받아들이는 특징인 Log scale 연산과정을 거친 벡터 값에 대한 주파수 특성을 다시 추출하는 DCT 변환을 적용한다. 최종적으로 앞선 과정은 데이터 값이 13차가 되는 벡터로 변환이 되는데, 여기서 시간에 따른 벡터 값의 변화 정도를 측정하여 다시 벡터 값으로 가지게 되면

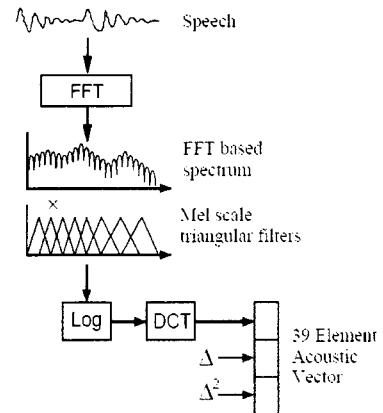


그림 4. MFCC feature extraction 과정

39차 특징 벡터가 되는 데이터 변환 과정이 마무리된다. 전체과정은 아래의 그림으로 요약 될 수 있는데, 여기서는 연산량을 줄이기 위해 13 벡터까지만 이용한 두 음성 데이터 값의 유사성을 계산하였다.

특징 추출 과정은 몇 가지 다양한 방법이 적용될 수 있으나, 인식 성능이 좋은 MFCC 방법이 주로 많이 사용되고 있다.

▪ **Dynamic Time Warping** : 이와 같이 특징이 추출된 데이터를 가지고 두 데이터 사이의 유사성을 계산하는 바문맥 종속 알고리즘 중 전통적으로 가장 많이 사용하는 Dynamic Time Warping (DTW)[1][2]이라는 방법을 사용하였다. 이것은 같은 발음이라도 말하는 사람에 따라 그 길이가 달라지는 문제가 있기 때문에 이를 해결하기 위해 사용하는 방법으로, 입력으로 들어오는 테스트 특징패턴의 시간 축 벡터열과 저장된 음성단위들이 대표패턴의 시간축 벡터열을 직접 비교하여 서로 대응되는 프레임을 찾아낸다. 이 경우 발음 속도의 차이에 따른 영향을 보상해주기 위해 시간 축 정규화 기술이 사용된다. 보통 하나의 단어를 발성하는 경우에 화자, 감정, 주변 환경에 따라 각각 다른 지속시간을 가지므로, 이러한 지속길이의

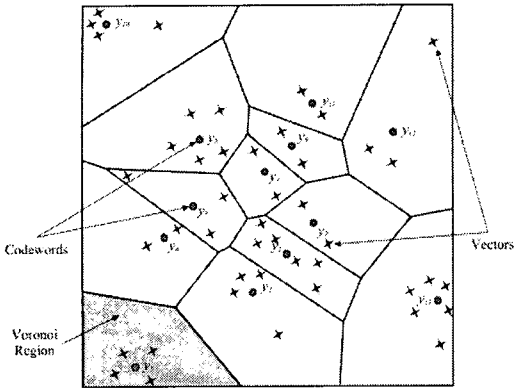


그림 5. 벡터 공간의 코드북

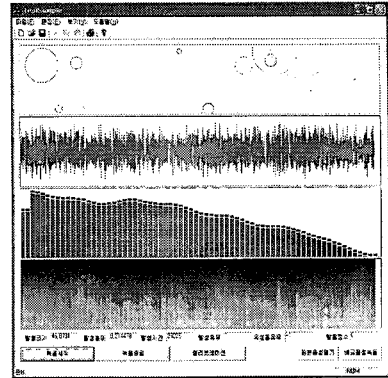


그림 6. 발성의 시각화

불일치를 비선형적으로 최적화하여 수행하는 특성을 갖는다. DTW는 인식 대상 어휘가 작은 독립 단어 인식에 주로 이용된다.

▪ **Kmeans알고리즘을 이용한 코드북 생성** : DTW는 두 벡터 사이의 거리를 계산하지만, 같은 발음에 대한 화자간의 차이를 잘 설명하지 못한다. 그래서 표준 발음 데이터에 대한 코드북을 생성하여 다른 화자와 표준 발음으로 정의된 화자간의 코드북 거리값을 계산한다. 이때 코드북을 만들기 위해 kmeans알고리즘을 적용하는데, 이것은 Vector Quantization방법의 한가지로 표준 발음 데이터인 벡터들의 분포를 코드워드(codeword) 중심으로 그룹화하여 1024개로 구분 짓는다. 아래 그림과 같이 벡터들은 가장 가까운 코드워드가 있는 구역으로 그룹 지어지게 된다.

화자간 발음 특성이 다를수록 코드북의 개별 그룹 중심인 코드워드와의 거리 값이 커지며, 반대로 발음 특성이 유사한 화자는 코드북의 코드워드와의 거리 값이 작아진다. 이와 같이 코드북 거리 값을 화자간 차이를 정규화시키는 방법으로 이용하여 DTW를 적용한 표준 발음과의 유사성을 계산한다.

본 연구에서 개발된 음성분석 기반 발성의 시각화 기술을 사용한 예제 S/W 결과는 다음과 같다.

### 3.2 시각과 음성 동기화 및 실감 처리 기술

본 연구에서는 대사 내용이 담긴 음성 파일(Wave File)과 대사 내용이 글로 저장된 텍스트 파일(Text File)을 입력으로 받아서 자동으로 립싱크 애니메이션 키프레임을 생성하는 기술개발을 최종 목표로 했다. 음성 기반 자동 립싱크 애니메이션 기술개발에서는 다음의 3가지 개발 내용이 핵심적인 부분을 차지하며 이 3가지 과정을 거쳐 자동립싱크 애니메이션 키프레임 데이터가 최종 결과물로 생성된다.

▪ **한글 발음을 가장 잘 표현할 수 있는 음소 모델을 정의하고 제작하는 기술** : 구화 애니메이션에 사용되어 한국어 발음입술 모양에 대한 시뮬레이션 기본 음소 모델을 정의하고 제작하는 과정이다. 발음음소 모델의 제작은 [3][4][5]을 참고해 정의하고 제작하였다.

한국어의 발음 구성은 다음과 같다.

- 자음 : ㄱ ㅋ ㆁ ㄷ ㅌ ㄹ ㄴ ㄷ ㅂ ㅃ ㅅ ㅆ ㅇ ㅈ ㅊ ㅊ ㅋ ㅌ ㅍ ㅎ (19자)

- 모음 : ㅏ ㅑ ㅓ ㅕ ㅗ ㅛ ㅜ ㅠ ㅡ ㅣ (21자)  
ㅝ ㅟ ㅛ ㅜ ㅞ ㅟ ㅡ ㅡ ㅣ (21자)

위의 발음을 입술, 혀, 목구멍의 위치와 모양 변화를 중심으로 총 30가지로 분류, 정의해 음소 발음 모델을 제작했으며 그 결과는 다음과 같다.



그림 7. 한국어 발음 음소 모델 제작 결과

■ 음성 분석을 통한 각 음소별 발음 시간 정보를 추출하는 기술 : 음성 분석을 통해 각 음소별 발음 시간 정보를 추출하는 기술이다. 발음이 녹음된 음성데이터와 음성 데이터의 발음 내용이 담겨있는 텍스트 정보를 입력으로 받아서 발음을 구성하는 각 음소의 발음 시작시간, 끝시간, 발음 유지 시간, 발음 세기 정보를 분석해 출력하는 기술을 개발하였다. 각 음소별 발음 시간의 분석은 HMM 모델 기반 음성 모델링과 음성 인식 및 분석 기술을 적용해 개발하였다.

분석된 발음 시간 정보는 자체 포맷인 SXT 파일정보로 출력되고 저장된다.

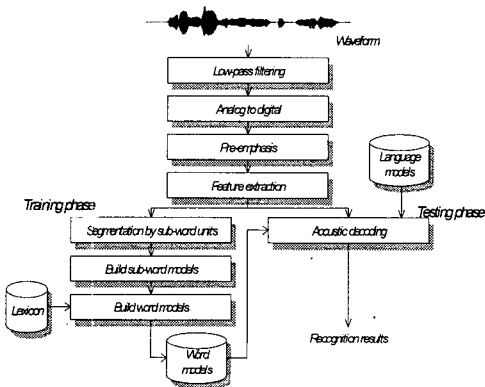


그림 8. 음성 분석 개요도

# SXT #		# SXT #	
# Ver. 2.0 #		# Ver. 2.0 #	
# Len 9 180 #	4888664.000	# Len 11.868 #	3513746.750
0.390 0.560 어		0.630 으	3513746.750
0.390 0.560 이		0.070 0.600 ㅏ	
0.590 0.730 언	39759000.000	0.600 0.630 ㅡ	
0.560 0.630 으		0.630 0.740 거	49770344.000
0.630 0.690 ㅓ		0.630 0.700 ㅑ	
0.690 0.790 ㅓ		0.700 0.740 ㅓ	
0.760 0.920 ㅓ	59041444.000	0.740 1.110 ㅓ	35726128.000
0.760 0.820 ㅓ		0.740 0.860 ㅓ	
0.820 0.850 ㅓ		0.860 0.890 ㅡ	
0.850 0.920 ㅓ		0.890 1.110 ㅓ	
0.920 1.130 ㅓ	22178936.000	1.180 1.260 ㅓ	30724906.000
0.920 1.010 ㅓ		1.180 1.260 ㅓ	
1.010 1.040 ㅓ		1.260 1.470 리	57023040.000
1.040 1.130 ㅓ		1.260 1.310 ㅓ	
1.130 1.190 ㅓ	19846270.000	1.310 1.470 ㅓ	
1.130 1.190 ㅓ		1.470 1.500 ㅓ	57784776.000
1.190 1.480 ㅓ	13563413.000	1.470 1.500 ㅓ	
1.190 1.330 ㅓ		1.530 1.710 사	4993568.000
1.330 1.480 ㅓ		1.530 1.650 ㅓ	
1.510 1.990 ㅓ	7804321.500	1.650 1.710 ㅓ	
1.510 1.890 ㅓ		1.710 1.810 ㅓ	40441772.000
1.890 1.990 ㅓ		1.710 1.740 ㅓ	
1.990 2.170 ㅓ	20539090.000	1.740 1.810 ㅓ	
1.990 2.070 ㅓ		1.840 2.030 ㅓ	15569140.000
2.070 2.120 ㅓ		1.840 1.940 ㅓ	
2.120 2.170 ㅓ		1.940 1.970 ㅓ	
2.200 2.320 ㅓ	22153634.000	1.970 2.030 ㅓ	
2.200 2.260 ㅓ		2.030 2.130 ㅓ	26639758.000
2.260 2.290 ㅓ		2.030 2.060 ㅓ	
2.290 2.320 ㅓ		2.060 2.130 ㅓ	
2.320 2.450 ㅓ	63007924.000	2.130 2.420 ㅓ	20499912.000
2.320 2.390 ㅓ		2.130 2.160 ㅓ	
2.390 2.450 ㅓ		2.160 2.420 ㅓ	
2.450 2.670 리	33105930.000	2.480 2.610 ㅓ	54164952.000
2.450 2.480 ㅓ		2.480 2.610 ㅓ	
2.480 2.670 ㅓ		2.510 2.750 리	33939672.000
2.700 2.940 ㅓ	35072668.000	2.510 2.690 ㅓ	
2.700 2.820 ㅓ		2.690 2.750 ㅓ	
2.820 2.850 ㅓ		2.780 2.890 ㅓ	24498694.000
2.850 2.940 ㅓ		2.780 2.890 ㅓ	

그림 9. SXT 정보

■ 음성 파일의 발음 시간 정보를 기반으로 해서 자연스러운 립싱크 애니메이션 키프레임을 자동으로 생성해 내는 기술 : 앞에서 생성된 SXT정보를 이용해 다음 과정을 거쳐서 음성에 동기화된 립싱크 애니메이션 키프레임 정보를 자동으로 생성한다.

- SXT 정보를 기본 애니메이션 그래프로 변환

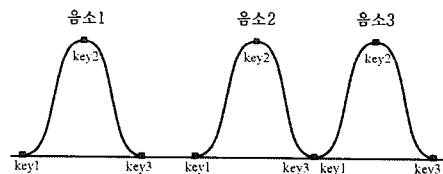


그림 10. 음소 Key Frame Setting

- 기본 애니메이션 그래프 수정 : SXT 정보에 포함된 발음 세기 정보를 Animation Curve에 적용하고 동시에 음소간 Blending 알고리즘도 적용함으로써 [음소-음소], [음절-음절] 의 입술 모양이 자연스럽게 애니메이션 될 수 있도록 Animation Curve를 수정한다.

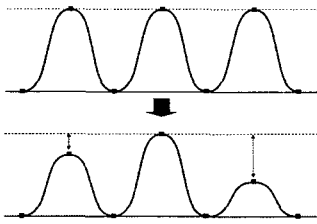


그림 11. 발음 크기 적용



그림 12. 음소간 Blending 적용

- 빠른 발음을 부드럽게 수정 : 발음 속도가 빠른 부분은 애니메이션이 자연스럽게 못하고 거칠게 표현될 수 있으므로 다음 두 개의 그림과 같은 강제적인 방법을 적용해 자연스럽게 부드러운 움직임을 얻을 수 있도록 수정과정을 거친다.

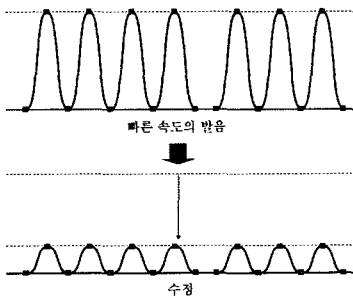


그림 13. 빠른 발음 수정 1

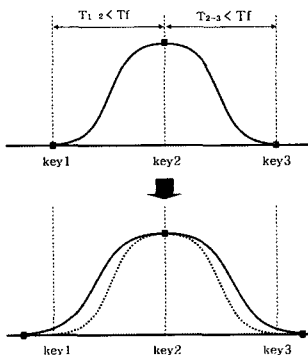


그림 14. 빠른 발음 수정 2

- 최종 키-프레임 생성 : 같은 모프-타겟의 Key Frame 이 겹칠 경우 그 부분의 애니메이션 커브를 다음 그림과 같이 최종적으로 수정한 후에, 수정된 그래프 상에서 다시 한 번 키-프레임을 추출함으로써 최종 완성된 립싱크 애니메이션 키-프레임 정보를 생성한다.

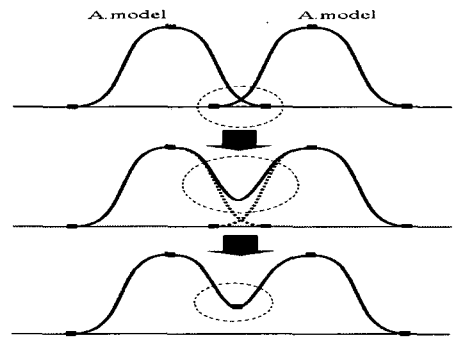


그림 15. 최종 키프레임 생성

본 연구에서 개발된 음성분석 기반 자동 립싱크 애니메이션 기술을 적용한 립싱크 애니메이션 S/W 및 결과 영상 이미지는 다음과 같다.



그림 16. 자동 립싱크 애니메이션 S/W

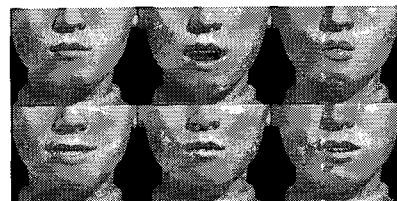


그림 17. 립싱크 동영상 이미지 예

### 3.3 우리말 학습 툴 개발

청각장애아동의 우리말 구화 교육에 사용될 수 있는 학습 툴은 다음과 같이 제작되었다.

- 생활 단어 1000개에 대한 음성 DB 구축
  - 자동 립싱크 애니메이션 기술을 적용해 이 음성에 동기화된 립싱크 애니메이션 동영상 을 제작
  - 립싱크 동영상을 통해 발음 소리와 발음시 입술의 모양 학습
  - 음성의 시각화 기술을 이용해 마이크를 통해 입력되는 청각장애아동의 발성 및 발음을 분석
  - 분석된 결과를 화면에 애니메이션으로 피드백
- 청각장애아동은 자신의 발성 및 발음을 시각적으로 확인하면서 학습 진행

음성분석을 통한 발성의 시각화 기술과 시각과 음성 동기화 및 실감 처리 기술을 적용한 우리말 학습 툴의 개발 결과는 다음과 같다.

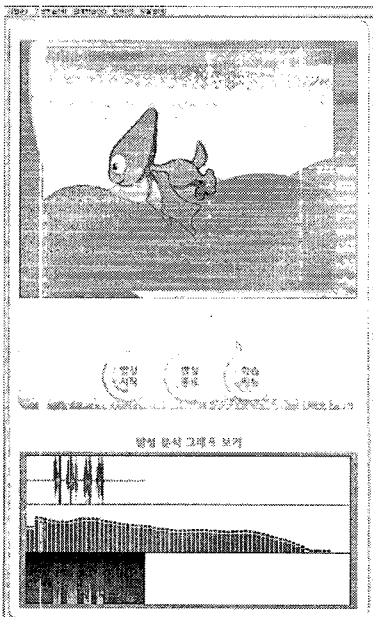


그림 18. 우리말 학습툴

### 3.4 구화교육용 애니메이션 콘텐츠 개발

본 연구에서는 “호두까기 인형”을 10분 정도의 분량으로 어린이들의 눈높이에 맞게 각색하고 본 연구에서 개발된 “음성 기반 자동 립싱크 애니메이션 기술”을 적용해 콘텐츠에 등장하는 3D 캐릭터의 립싱크 애니메이션에 적용함으로써 “2D + 3D 합성 애니메이션” 제작 방식을 통해 청각 장애를 가진 아동들과 일반 아동들이 함께 즐길 수 있는 애니메이션 콘텐츠를 제작하였다. 애니메이션의 제작 과정은 다음과 같다.

- 시나리오 작업
- 성우 녹음 작업
- 설정 작업
- 콘티 작업
- 3D 캐릭터 설정 및 모델링, 세팅 작업
- 레이아웃 작업
- 배경 원도 제작
- 원화 제작
- 동화 제작
- 채색 작업
- 3D 립싱크 애니메이션, 모션 애니메이션 제작
- 배경 제작
- 2D + 3D 합성 및 애니메이션 완성

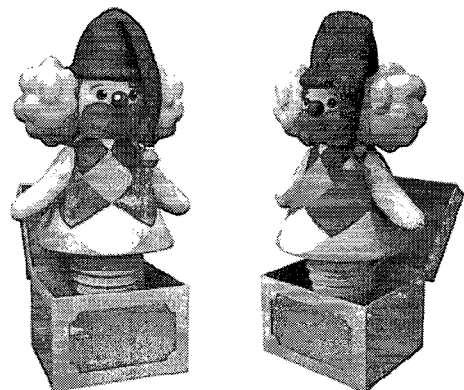


그림 19. 애니메이션 콘텐츠 예



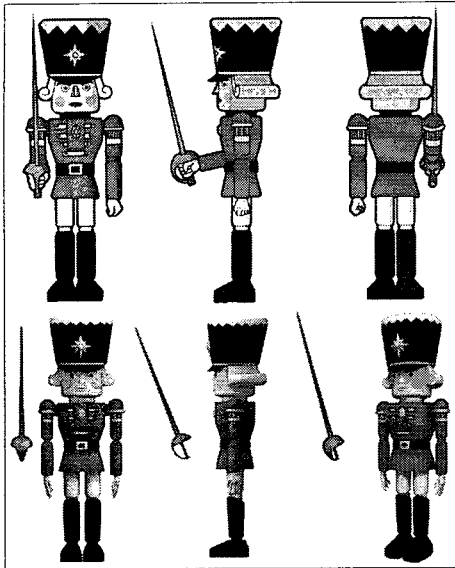


그림 20. 애니메이션 콘텐츠 예



그림 21. 애니메이션 콘텐츠 예

#### 4. 결론

본 논문에서는 청각장애아동의 구화교육을 위한 애니메이션 콘텐츠 제작기술에 필요한 두 가지 핵심적인 기술의 개발 내용 및 개발 결과를 살펴보고, 그 두 가지 기술을 활용한 결과물인 우리말 학습 툴과 콘텐츠 제작에 관한 결과를 간략하게

살펴보았다.

아직까지 순수 국내 기술을 이용한 청각장애인들의 발음 학습을 돕는 제품은 국내에서 출시되지 않았으며, 외국 기술을 사용한 제품만이 나와 있는 상황에서, 본 연구에서 개발한 기술이 상용화된다면 해외 기술에 대한 로열티 지급을 줄일 수 있고, 더불어 수입 대체 효과를 기대할 수 있으며, 청각 장애인들을 위한 디지털 콘텐츠 및 기술 개발을 촉진시키는 계기가 될 수 있을 것으로 기대된다. 또한 단순한 2D 애니메이션을 활용한 놀이 수준의 기존 발음 학습 툴들의 한계를 넘어서, 실제 발음 내부 구조, 발음 입술 모양, 근육의 움직임, 호흡의 변화 등을 3D로 정확하게 재현해 제공함으로써 효율적이고 정확한 발음 학습을 필요로 하는 청각 장애인 및 청각 장애인 교육 기관에서 다양한 용도로 사용 가능할 것이다.

장애인을 위한 학습 툴을 제작하는 것 이외에도 3D CG 업계에서 사용되고 있는 음성 데이터를 이용한 립싱크 애니메이션 저작 도구 역시 외국 기술을 사용한 저작 툴들은 많이 나와 있지만, 순수 국내 기술로 개발된 저작 툴은 아직 개발되어 있지 않으므로, 본 과제에서 개발하는 립싱크 애니메이션 기술은 해외 소프트웨어 사용을 대체할 수 있을 것으로 기대되며, 3D CG 업체에서 실제로 사용될 경우 애니메이션 제작에서 상당한 시간과 비용의 절감 및 결과물의 질적 향상을 가져올 수 있을 것이다.

#### 참고 문헌

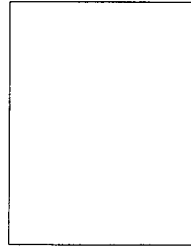
- [1] C. Myers, L. Rabiner, and A. Rosenberg, "Performance Tradeoffs in Dynamic Time Warping Algorithms for Isolated Word Recognition", IEEE Trans. on acoustics, speech, and signal processing, vol. ASSP-28, No.6, Dec 1980.

[2] P.H.W. Wong, Oscar C. Au, J.W.C. Wong, W.H.P. Lau, "Reducing Computational Complexity of Dynamic Time Warping Based Isolated Word Recognition with Time Scale Modification," in Proc. of Fourth Int. Conf. on Signal Processing (ICSP), vol. 1, pp.722-725, Beijing, Oct. 1998.

[3] Frederic I.Parke and Keith Waters, "Computer Facial Animation"

[4] 김승국 "발성-발화 지도의 이론과 실제" 특수학교 교육과정 운영자료 11

[5] 고도홍, "언어기관의 해부와 생리 - Anatomy & Physiology of Speech Organs"



오 범 수

- 1996년 고려대학교 컴퓨터공학과 (공학사)
- 1998년 고려대학교 대학원 컴퓨터공학과 (공학석사)
- 2002년 고려대학교 대학원 컴퓨터공학과 (공학박사)
- 2002년 3월 ~ 현재 비쥬텍쓰리디(주) 대표이사
- 관심분야 : 3D 컴퓨터그래픽스, 2D 컴퓨터그래픽스, 프로그래밍언어, 멀티미디어 시스템



조 성 업

- 1997년 고려대학교 전자공학과(공학사)
- 1999년 고려대학교 대학원 전자공학과(공학석사)
- 2002년 고려대학교 대학원 전자공학과 (공학박사 수료)
- 2002년 2월 ~ 현재 비쥬텍쓰리디(주) 책임연구원
- 관심분야 : 3D 컴퓨터그래픽스, 2D 컴퓨터그래픽스, 시뮬레이션



임 철 수

- 1985년 서울대학교 계산통계학과 (학사)
- 1988년 Indiana University(미) 전산과학과 (석사)
- 1994년 서강대학교 전자계산학과 (박사)
- 1985~1996년 (주)데이콤, (주)아시아나항공, (주)신세기통신 근무
- 1997년 3월 ~ 현재 서경대학교 컴퓨터공학과 교수
- 관심분야 : 멀티미디어시스템, 차세대컴퓨팅 응용 기술